A Survey of 10 Gigabit Ethernet and Backplane Ethernet

Jacob Li , lizihao@wustl.edu (A paper written under the guidance of Prof. Raj Jain) Download

Abstract

Ethernet is growing rapidly now, and its transmission rate is increased from 1000 bps to 100 Gbps and even 400 Gbps. Higher-speed Ethernet is on the way, and problems are followed. This paper introduces the procedure of high-speed Ethernet's development and Backplane's emergence and discusses issues of every development stage like interoperability between devices. The paper also discusses how technologies like Serializer/Deserializer (SerDes) and Auto-Negotiation solves those problems in practice and push the progress of high-speed Ethernet development.

Keyword: Backplane Ethernet, 10 Gigabit Ethernet, SerDes, Auto-Negotiation, IEEE 802.3ap, 10GBASE, 40GBASE, Physical Layer

Table of Contents:

1. Introduction

In networking, Ethernet plays a very important role and is widely used to provide connectivity within and between systems. Now, as Ethernet is growing faster, the Ethernet can even provide faster links over multiple media [McCoubrey16]. As the demand for high-speed communication technology increases, the Ethernet's deployment all over the world results in bandwidth requirement explosion. The bandwidth has already reached 10/40 Gbps and stepping forward to higher speed. The next-generation serial operation interfaces between 50 Gbps and 60 Gbps are on their way to deploy 400 Gbps Ethernet systems. High-speed buses standards like 10 Gigabit Attachment Unit Interface (XAUI), Interlaken, and System Packet Interface version 4.1/4.2

(SPI4.1/SPI4.2) can meet the 10/40 Gbps bandwidth requirement, but their transmission distances are limited within 50 centimeters [IEEE01]. Thus, for a high-performance computing platform, long transmission distance is required. To solve this problem, the IEEE published the IEEE 802.3ap standard and included the concept of Backplane Ethernet, providing us with specifications like 10GBASE-KR and 40GBASE-KR4, whose maximum transmission distance are 1 meter.

So, what is Backplane Ethernet? Basically, the Backplane Ethernet is Backplane 10 Gigabit Ethernet and mostly applied to backplane applications like blade servers, routers, switches, and some embedded devices now [Held13]. It uses existing Ethernet standard to build up an optimal interface for backplane. It will not define areas like mechanical or material but focuses on electrical and channel model specifications [Ghiasi05].

Figure 1 shows that Backplane Ethernet specifications like 10GBASE-X, 10GBASE-BX, 10GBASE-KX, and 10GBASE-KR/KR4 represent high-speed transmission in each time and the link speed keeps increasing as new specification comes up [McCoubrey16]. The Backplane Ethernet requires the equipment producers to build up high-speed networking backplane for servers using off the shelf or merchant silicon, which means that the manufacturers have to build something very similar to the high-performance virtual connect but only to use Ethernet [Ferro11].

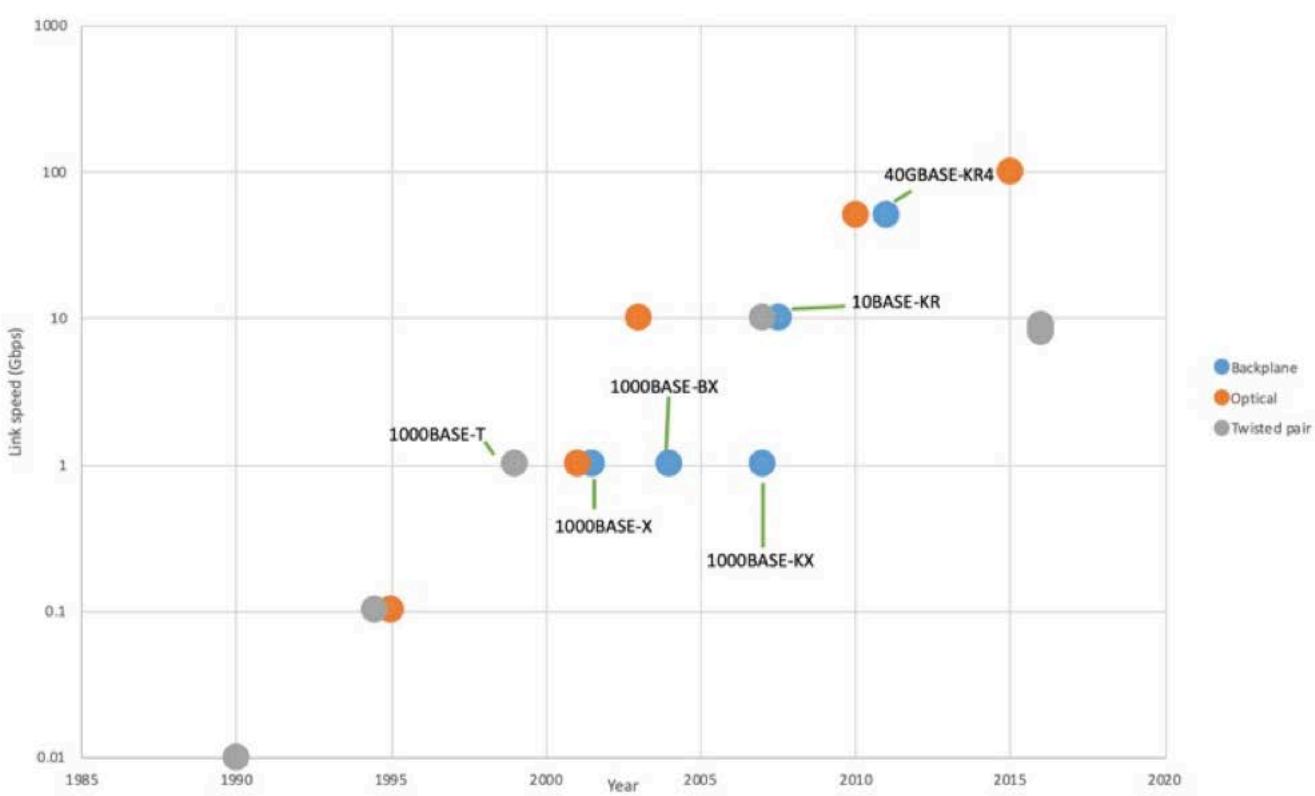

Figure 1: Ethernet standards development in link speed

In 2007, 802.3ap was published, and 802.3 Ethernet had a formal Physical Layer (PHY) standard for electrical backplanes formally. Backplane Ethernet was included in the 802.3ap. The benefits of initial Serializer/Deserializer (SerDes) Ethernet are offered by 1000BASE-KX. For example, Gigabit performance over two differential pairs across electrical backplanes [McCoubrey16].

The later addition of the 802.3 standard included the Backplane Ethernet PHY standards for 10 Gbps, 40 Gbps, and 100 Gbps. 10 Gigabit Ethernet over one single lane and 40 Gigabit Ethernet over four lanes came true, supported by 10 Gbps SerDes technology in the systems with high-performance modules and backplanes. All the 802.3 Backplane Ethernet standards have the feature of auto-negotiation. Thus the connection between devices with different native PHY types is possible [McCoubrey16].

The Backplane Ethernet, together with relevant 10 Gigabit Ethernet standards and specifications, higher-speed and longer-distance transmissions become true and more cutting-edge devices and applications using Backplane Ethernet are on their way.

2. 10 Gigabit Ethernet Standards

In the procedure standardizing of Gigabit Ethernet, IEEE and 10 Gigabit Ethernet Alliance (10 GEA) are two of the most important organizations. There are multiple different Gigabit Ethernet standards and keep extending as demands increase, and technology develops. First, IEEE 802.3ae was published in 2002. Then continuing standards came up, like IEEE 802.3an standard for 10GBASE-T over unshielded twisted pair (UTP) and IEEE 802.3aq standard for 10GBASE-LRM over optical fiber for 10 Gbps Ethernet in 2006; IEEE 802.3ap standard for Backplane Ethernet over printed circuit boards in 2007. Then, standards for higher-speed and longer-distance transmission followed. In 2010, IEEE 802.3ba was published for 40/100 Gbps Ethernet, in which 40 Gbps transferred over 1-meter backplane and 100 Gbps up to 10 meters. After that, IEEE 802.3bm came up in 2015 for 100/40G Ethernet over optical fiber. 802.3cb was published for 2.5/5 Gbps operation over backplane in 2018 [IEEE02].

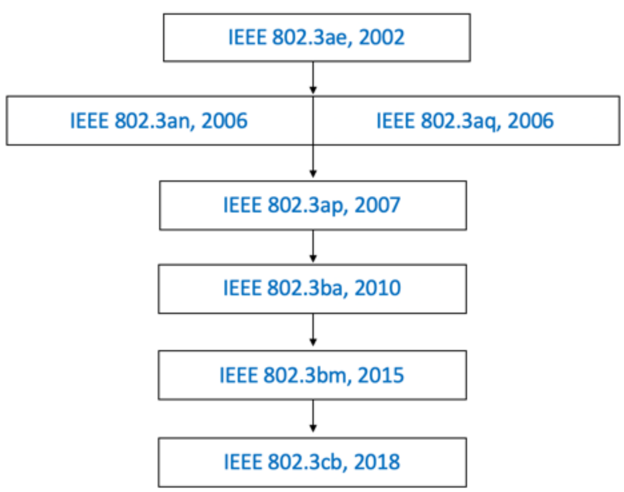

Figure 2: Selected 802.3 Standards Development Timeline

2.1 IEEE 802.3ae 10 Gigabit Ethernet PHY Structure

IEEE 802.3ae is the beginning of the 10 Gigabit Ethernet standard. There are multiple standards and specifications defined in it. The IEEE 802.3ae standards includes three physical interface standards, they are 10GBASE-R (10GBASE-SR, 10GBASE-LR and 10GBASE-ER), 10GBASE-W (10GBASE-SW, 10GBASE-LW and 10GBASE-EW) and 10GBASE-X (only includes 10GBASE-LX4). Figure 3 shows the 10GBASE-X structure. The physical layer structure is basically similar to the Gigabit Ethernet, only the interface between Physical Coding Sublayer (PCS) and Reconciliation Sublayer (RS) is 10 Gigabit Media Independent Interface (XGMII) instead of the previous Gigabit Media Independent Interface (GMII). Figure 4 shows the 10GBASE-R structure; besides the XGMII interface, another difference is the coding scheme changed from 8B/10B to 64B/66B [IEEE02].

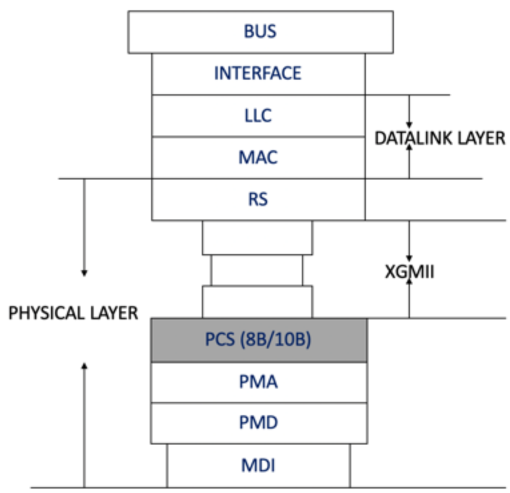

Figure 3: 10GBASE-X PHY Structure

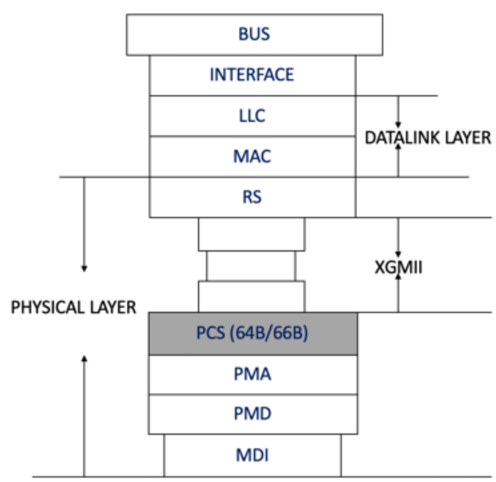

Figure 4: 10GBASE-R PHY Structure

10GBASE-W has the most modification in its physical layer, comparing with the other two standards. Figure 5 shows the structure of 10GBASE-W: another obvious change in the 10GBASE-W is adding a new layer between PCS and Physical Media Attachment (PMA), the Wide Area Network Interface Sublayer (WIS). The 10 Gigabit Ethernet can be switched to low transmission rate by WIS, which allows 10 Gigabit Ethernet devices be compatible with Synchronous Optical Network (SONET) [IEEE02].

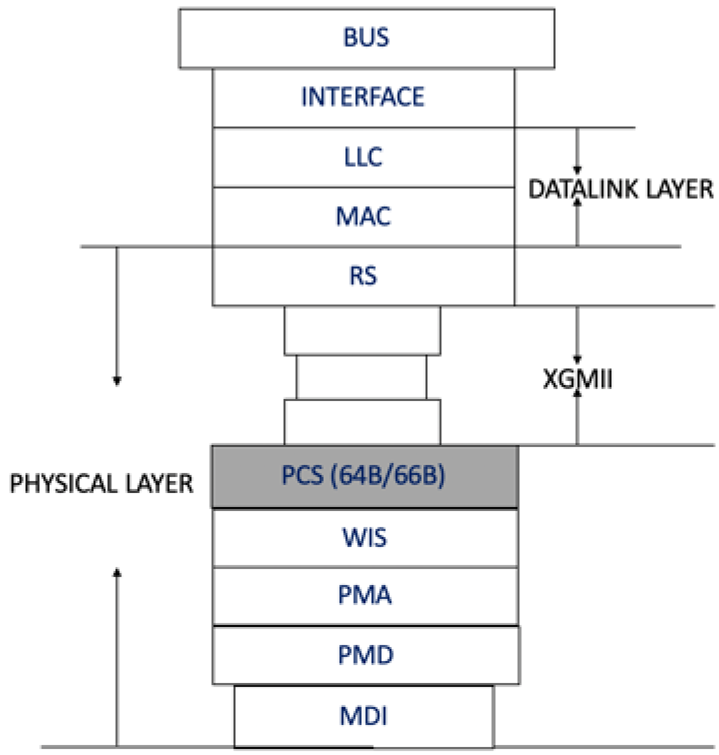

Figure 5: 10GBASE-W PHY Structure

2.2 10 Gigabit LAN Ethernet over Optical Fiber

For now, the 10 Gigabit LAN (Local Area Network) Ethernet specifications include: 10 GBASE-SR, 10GBASE-LR, 10GBASE-LRM, 10GBASE-ER, 10GBASE-ZR and 10GBASE-LX4, Figure 6 [IEEE02].

The SR in the 10GBASE-SR means Short Reach. This specification supports shortwave (wavelength is 850 nanometers) Multiple-Mode Fiber (MMF), whose coding scheme is 64B/66B, the transmission distance is from 2 to 300 meters. The 10GBASE-SR has the advantages of low price, minimum power cost and smallest size of fiber block [IEEE02].

The LR in the 10GBASE-LR means Long Reach. This specification supports longwave (wavelength is 1310 nanometers) Single-Mode Fiber (SMF) whose coding scheme is 64B/66B, the transmission distance is from 2 meters to 10 kilometers. In practice, the transmission can reach 25 kilometers at max. The price of 10GBASE-LR optical fiber block is lower than 10GBASE-LX4 optical fiber block [IEEE02].

The ER in the 10GBASE-ER means Extended Reach. This specification supports super longwave (wavelength is 1550 nanometers) SMF. The transmission distance is from 2 meters to 40 kilometers [IEEE02].

The 10GBASE-LX4 takes wavelength-division multiplexing technology, using 4 lanes whose wavelengths are all 1300 nanometers. The transmission distance of this specification in SMF can reach 10 kilometers. It mainly used in the situation that the same optical fiber block supports SMF and MMF at the same time [IEEE02].

The LRM in 10GBASE-LRM means Long Reach Multimode, corresponding to the IEEE 802.3ap standard that came up in 2006. The wavelength is 1300 nanometers. The transmission distance can reach 260 meters in OM3, which is a common multimode optical fiber [IEEE02].

Several manufacturers came up with a module interface whose transmission distance reaches 80 kilometers. This is the 10GBASE-ZR specification. It uses a super long wavelength (1550 nanometers) SMF. However, the 80 kilometers physical layer is not in the IEEE 802.3ar standard, it was described in Optical Carrier (OC) -192/ Synchronous Transport Module (STM) -64 Synchronous Digital Hierarchy (SDH)/SONET specification by manufacturers, so it will not be accepted by IEEE [IEEE02].

2.3 10 Gigabit WAN Ethernet over Optical Fiber

The 10GBASE-SW, 10GBASE-LW, 10GBASE-EW, and 10GBASE-ZW specifications are physical specifications applied to the Wide Area Network (WAN). They are set to work in OC-192/STM-64 SDH/SONET. The optical fiber types and transmission distance correspond to the 10GBASE-SR, 10GBASE-LR, 10GBASE-ER and 10GABSE-ZR specifications introduced in section 3.2. 10GBASE-LX4 and 10GBASE-CX4 specifications have no WAN physical layer because the SDH/SONET standard works in serial transmission, but in the 10GBASE-LX4 and 10GBASE-CX4 specifications use parallel transmission [IEEE02].

2.4 10 Gigabit Ethernet over Twisted-Pair Cable

10 Gigabit Ethernet based on double-twisted cable includes 10GBASE-CX4, 10GBASE-KX4, 10GBASE-KR, and 10GBASE-T [IEEE02].

IEEE 802.3ak came up with 10GBASE-CX4 for 10 Gigabit Ethernet transmission by coaxial cable and was approved in 2004. However, the transmission distance is limited to 15 meters. 10GBASE-CX4 uses XAUI defined in the 802.3ae and 4X connector used in InfiniBand. The transmission media is called CX4 copper cable [IEEE02].

10GBASE-KX4 and 10GBASE-KR correspond to the IEEE 802.3ap standard published in 2007. They are mainly used for backplane applications, like blade servers and routers so they also as known as Backplane Ethernet. 10 Gigabit backplane has serial version and parallel version now. The parallel version (10GBASE-KX4) is the general design of the backplane, it separates the 10 Gigabit signal into four lanes, like XAUI. 10GBASE-KX4 uses the same physical layer coding like 10GBASE-CX4; the bandwidth of every lane is 3.125 Gbps. In the serial version

(10GBASE-KR), only one lane was defined. It uses the same physical layer coding with 10GBASE-LR/ER/SR, the 64B/66B coding scheme to realize 10 Gbps high-speed transmission [IEEE02].

3. SerDes

Before Backplane Ethernet, the mainly used specification is 1000BASE-T, which was defined in IEEE 802.3ab in 1999. 1000BASE-T offered up to 100 meters Gigabit Ethernet of unshielded twisted pair cable, which is widely used in many commercial applications. 1000BASE-T is very popular because it combines high-performance, robustness, and inter-operability. However, the 1000BASE-T has some shortcomings. First, bulky transformers are needed to provide magnetic coupling. Also, it may need a discrete physical layer (PHY) or controller device, and it needs four differential pairs with two send, two receive. These features may cause the size, weight and power issues, especially when these are critical concerns [McCoubrey16].

To solve these issues, SerDes Ethernet was introduced. SerDes refers to the Serializer/Deserializer chips that perform the physical signal generation and was originally used for chip-to-chip communication. It only needs two signal pairs, one for sending and the other for receiving. SerDes is a pair of functional blocks that are widely used in high-speed data transmission to compensate for limited input/output. The SerDes consists of two blocks: Parallel In Serial Out (PISO) block and Serial In Parallel Out (SIPO) block. These blocks convert data between serial data and parallel interfaces in each direction. There are four different SerDes architectures: 1. Parallel clock SerDes 2. Embedded clock SerDes 3. 8b/10b SerDes 4. Bit interleaved SerDes [Lewis04].

There are several standards defined for SerDes Ethernet. In table 1, there are four selected standards for SerDes Ethernet. 1000BASE-X, 1000BASE-BX, 1000BASE-KX, and Serial Gigabit Media Independent Interface (SGMII) have same maximum data transmission rate, which is 1 Gbps and 10GBASE-KR has data transmission rate of 10 Gbps at max, which was defined in IEEE 802.3ap. Another difference between 10GBASE-KR and others is its line coding way. It uses 64B/66B instead of 8B/10B that others use. Other than standards listed in Table 1, there are also standards like 1000BASE-LX and 1000BASE-SX, all of these share the same encoding and signaling layer, which is referred to as "BASE-X." The SGMII is a specification that provides connections between separate MAC and PHY devices. These devices leverage a single SerDes pair at Gigabit rates with "BASE-X" encoding, too [McCoubrey16]. The 1000BASE-KX has the features of auto-negotiation and capability advertisement, just like other 802.3 standards. This matches the purpose of making devices with different capabilities or underlying techniques work together properly, which greatly simplifies the interoperability [McCoubrey16].

Table 1: Selected SerDes Ethernet Standards

Standard	Data Rate	Symbol Rate	Line Coding	Standardized
1000BASE-X	1 Gbps	1.25 Gbaud	8B/10B	IEEE 802.3z
1000BASE-BX	1 Gbps	1.25 Gbaud	8B/10B	PICMG ATCA 3.1
1000BASE-KX	1 Gbps	1.25 Gbaud	8B/10B	IEEE 802.3ap
SGMII	1 Gbps	625 MHz DDR	8B/10B	-
10GBASE-KR	10 Gbps	10.3125 GHz	64B/66B	IEEE 802.3ap

Since those SerDes standards use the same encoding, devices based on them can be connected and work together. Moreover, the 1000BASE-BX specification also solves the problem of interoperability gap, which is caused by the specification difference. The SerDes Ethernet plays a very important role before Backplane Ethernet. It solves problems existing in the previous 1000BASE-T.

4. Auto-Negotiation

Ethernet's transmission speed is approaching 400 Gbps now. Auto-Negotiation is an algorithm defined in IEEE 802.3u to make fast Ethernet comes true, which is very important to the Ethernet architecture and technology. To make sure interoperability between systems and success in reaching 400 Gbps speed in the end, Auto-Negotiation has to be implemented.

Actually, the Auto-Negotiation procedure was designed in a very simple way. It is a handshaking algorithm that relies on devices advertising what kind of technologies are supported by setting particular bits. Each bit indicates a specific technology in a defined length. Then, to begin, link signaling uses a highest common denominator approach [Lapak16].

Although the Auto-Negotiation itself is simple, there is a problem that non-complaint devices and misunderstanding of how the standards were designed to work disable the Auto-Negotiation badly. The reason for the problem is that both the link speed and the duplex were configured manually before the Auto-Negotiation. During standardization, if the device to be connected did not support Auto-Negotiation, the Auto-Negotiation will detect it and then enables half-duplex. This led to massive packet loss. Luckily, when standardization of Gigabit Ethernet was done, Auto-Negotiation would establish links to take more thorough testing and compliance with higher rates [Lapak16].

As Backplane Ethernet came out, another issue followed. Although the Auto-Negotiation works well for Ethernet over structured cable (twisted pair), including speeds like 2.5, 5, 25, and 40GBASE-T, its underlying technology does not fit well in a high-speed serial-link-based environment, especially backplanes. To solve this problem, a new Auto-Negotiation protocol was defined by IEEE 802.3 Clause 73. The new protocol was based on the previous algorithm based on twisted-pair. The difference is the new one takes advantage of both high speed and continuous signaling. This was designed mainly for the backplane physical layer [Lapak16].

Since all the 802.3 Backplane Ethernet standards have the feature of Auto-Negotiation, the connection between devices with different native PHY types is possible. According to this, the devices will negotiate and choose to use the optimal connection, which is supported by them. Table 2 shows the scheme of choosing the optimal common type is defined in 802.3 clause 73, 100GBASE-CR10 with 100 Gbps transmission rate using 10 lanes has the highest priority. Then, 40GBASE-CR4/KR4 with 40 Gbps transmission rate using 4 lanes, 10GBASE-KR with 10 Gbps rate using 1 lane, 10GBASE-KX4 with 10 Gbps rate using 4 lane4. 1000BASE-KX has the lowest priority with low transmission rate, 1 Gbps, using 1 lane [IEEE01] [McCoubrey16].

Table 2: 802.3 Clause 73 Priority

Priority	Technology	Capability
1	100GBASE-CR10	100 Gbps 10 lane, highest priority
2	40GBASE-CR4	40 Gbps 4 lane
3	40GBASE-KR4	40 Gbps 4 lane
4	10GBASE-KR	10 Gbps 1 lane
5	10GBASE-KX4	10 Gbps 4 lane
6	1000BASE-KX	1 Gbps 1 lane, lowest priority

For all Ethernets whose speeds are over 10 Gbps, Auto-Negotiation conformance is very important. In addition to the optimal operating speed selection, Auto-Negotiation is now used as a precursor to link establishment's second stage, which is called "link training." The transmitters and receivers of devices will be configured by link training, which optimizes electrical performance through a mutual data exchange before the establishment of the link. This optimization is not possible when devices are connected manually [Lapak16].

High-speed Ethernet increases from 10 Gbps, 40 Gbps, and 100 Gbps to 200 Gbps and 400 Gbps, Auto-Negotiation technology plays such a vital role that it needs to be addressed as soon as possible to make interoperability and the success of these high speeds come true.

5. Summary

During this survey of Backplane Ethernet and 10 Gigabit Ethernet standards, we go through the whole development procedure of these standards and how these standards support the Ethernet rate at 10 Gbps, 40 Gbps, 100 Gbps, etc. In this procedure, we know how SerDes solves the previous 1000GBASE-T's shortcomings and plays an important role before Ethernet Backplane. Then, as Ethernet rate increases, Auto-Negotiation technology makes sure the success of high-speed Ethernet and interoperability between devices. Previous high-speed Ethernet standards and technologies like Backplane, SerDes, and Auto-Negotiation build up foundation of incoming Ethernet with higher speed and better performance.

6. References

1. [McCoubrey12] Andrew McCoubrey, Backplane Ethernet: Enabling High-Performance Embedded Systems, Curtiss-Wright Defense Solutions, 2016, https://www.eletimes.com/wp-content/uploads/2016/12/backplane-ethernet-enabling-high-performance-embedded-systems.whitepaperpdf.render_2.pdf
2. [IEEE01] IEEE, 802.3ap, "Part 3 Amendment 4: Ethernet Operation over Electrical Backplanes," 2007, https://ieeexplore.ieee.org/servlet/opac?punumber=4213274
3. [IEEE02] IEEE, 802.3, "IEEE Standard for Ethernet," 2015, https://ieeexplore.ieee.org/servlet/opac?punumber=7428774
4. [Held13] Gilbert Held, Windows Networking Tools, CRC Press, 2013, ISBN: 9781466511071, https://books.google.com/books?id=gDzSBQAAQBAJ
5. [Ghiasi05] Ali Ghiasi, IEEE 802.3ap "Backplane Ethernet" Overview, 2005, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.396.3479&rep=rep1&type=pdf
6. [Ferro11] Greg Ferro, Backplane Ethernet - the GbaseK Standard, 2011, https://etherealmind.com/backplane-ethernet-gbase-kr-kx/
7. [Ferro10] Greg Ferro, Futures Review on 40 and 100 Gigabit Ethernet, 2010, https://etherealmind.com/notes-physical-connectors-40-100-gigabit-ethernet/
8. [Lewis04] Dave Lewis, et al., SerDes Architectures and Applications, DesignCon, 2004, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.173.4371
9. [Lapak16] Jeffrey Lapak, Michael Klempa, The Importance of Auto-Negotiation in Ethernet Architectures, 2016, https://www.electronicdesign.com/ethernet/importance-auto-negotiation-ethernet-architectures

7. List of Acronyms

- LAN: Local Area Network

- MMF: Multiple-Mode Fiber
- PCS: Physical Coding Sublayer
- PHY: Physical Layer
- PMA: Physical Media Attachment
- RS: Reconciliation Sublayer
- SerDes: Serializer/ Deserializer
- SGMII: The Serial Gigabit Media Independent Interface
- SMF: Single-Mode Fiber
- SONET: Synchronous Optical Network
- SPI4.1/4.2: System Packet Interface version 4.1/4.2
- WAN: Wide Area Network
- WIS: Wide Area Network Interface Sublayer
- XAUI: 10 Gigabit Attachment Unit Interface
- XGMII: 10 Gigabit Media Independent Interface

Last Modified: December 10, 2019
This and other papers on latest advances in computer networking are available on line at
http://www.cse.wustl.edu/~jain/cse570-19/index.html
Back to Raj Jain's Home Page